Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 12:24:29
REFERENCE NO: 233

Author Names & Affiliations

- David Lampert - Oklahoma State University School of Civil and Environmental Engineering

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Contaminant fate and transport modeling, Hydrologic modeling, Wastewater treatment, Water and wastewater systems, Life cycle assessment

Title of Submission

Incorporating Advanced Computing and Cyberinfrastructure into Environmental and Water Resources Engineering

Abstract (maximum ~200 words).

Many scientific disciplines rely heavily on computing and cyberinfrastructure to simulate physical systems, store information, and analyze the relationships between observed variables. The environmental and water resources engineering community has been slow to adopt these approaches, however. Software is widely used, but much of it is commercial or closed source. Disconnects often exist between the underlying mathematics and software application. Many environmental models utilize data-driven algorithms based on older observational approaches. Limitations in these empirical models now prohibits progress in the discipline. The educational curriculum, perhaps due to its breadth, places little emphasis on advanced algorithm development or coding. There is a critical need for a new generation of computational environmental engineers to build and apply advanced cyberinfrastructure to solve environmental problems. Novel algorithms are needed to create reproducible models of watershed hydrology. Open source codes should be developed for simulating water quality and water treatment processes. Cyberinfrastructure is needed to support big data analytics to enhance the sustainability and security of water and energy systems. An educational workforce and a culture of open source software development would facilitate knowledge transfer, address challenges in maintenance of legacy codes, and increase model reproducibility with potentially transformative results.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Growth in the world's population and the associated demands for food, energy and water increasingly stress the planet's limited fresh water resources. Environmental and water resources engineers determine how to supply water of the appropriate quality and quantity for its many purposes while maintaining critical ecosystem services. Water management issues are intimately connected with computing and simulation tools. Computing resources are used to analyze quantities of groundwater and surface water resources, determine the fate and transport of pollutants, estimate water usage in engineered distribution systems, and optimize of water and wastewater treatment processes. Because of the ubiquitous nature and widespread utilization of water resources, there is tremendous need and opportunity for employment of advanced cyberinfrastructure to improve water security and sustainability.

While environmental and water resources engineers utilize many cyberinfrastructure tools, the adoption of new advanced computing and modern programming practices is limited. There are a number of possible explanations for this behavior. A common issue is the shortage of data. Members of the community often disregard the value of advanced computing methods due to data limitations: “we don’t have the data to justify building a more complex model.” An additional challenge may be the inherently public nature of environmental resources. Water and air resources are not subject to market forces where economic incentives exist to justify financial support for the collection of additional information. Another reason may be the breadth of the educational requirements—environmental and water resources engineering encompasses aspects of geosciences, chemical sciences, biological sciences, and fluid mechanics. Programs are forced to limit the coursework in the plans of study, and advanced computing is not typically prioritized. There is a critical need for adoption of novel algorithms, workflow practices, and open source models within the environmental and water resources engineering community. The following discussion highlights a few particular areas of need including hydrologic modeling, open life cycle assessment databases, water treatment process models, and water system analytics tools.

Hydrologic Modeling. Digital simulations of hydrology have long been used to forecast supplies of surface and groundwater resources. The Stanford Watershed Model was developed in 1966 by Crawford and Linsley [1] to forecast the hydrologic regimes of streams and rivers. Such models provide critical insight into water resources planning and management. During the 1960s, the hydrology group at Stanford used more computing time than all groups other than high-energy physics [2]. Since these initial developments, models have increased in complexity to account for the extreme variabilities in hydrologic regimes and multitude of relevant hydrological processes. Many federal and state agencies now have their own hydrologic modeling tools. Watershed models provide an accounting of all fluxes of water within a study area including precipitation, evapotranspiration, baseflow, surface runoff, and anthropogenic withdrawals. These models have many important applications including water supply estimation, pollution fate and transport analysis, and flood inundation modeling. Hydrologic models rely principally on land use and climate data to predict stream flow and water quality. A set of parameters are typically “calibrated” to fit the simulated runoff to the values observed at a gauging station (often a single point in space). Despite the advancement in computing efforts, forecasting of runoff in ungauged watersheds remains an unsolved challenge. Most models require some form of calibration, which introduces arbitrary decision-making by the modeler. The majority of publications in the hydrology literature focus on specific watersheds to study particular phenomena and water management issues. Groundwater and surface water are often modeled independently, despite their interconnected nature. Legacy codes are widely used that were derived from empirical algorithms that have little physical basis. There is resistance from communities to changing practices, however. Several recent publications have questioned the scientific nature of hydrological models due to their lack of reproducibility [3]–[6]. Advanced cyberinfrastructure is needed to develop reproducible hydrologic models that do not require calibration.

Open Life Cycle Assessment Databases. Sustainability analysis is often performed by environmental engineers through the development of life cycle assessments (LCA) for consumer products. LCA is an analytical approach for assessing the environmental impact of a product over its whole life cycle from raw material extraction to disposal or reuse. LCAs can provide the net greenhouse gas emissions, water consumption, human health effects, etc. for similar products. LCAs are critical for formulating appropriate science-based energy and environmental policies. Databases of mass and energy flow for unit process are essential for LCA, but the most widely used tools (SimaPro) are commercial.

Open Source Water Treatment Process Models. Environmental engineers design wastewater treatment plants, which consume 3% of all electricity generated in the US despite having the potential to act as net energy producers [7]. The design of wastewater treatment requires forecasting of many complex biochemical processes. However, there are few open source tools to simulate the behavior of such systems. Such tools could assist with optimization to enhance performance. Inorganic water speciation chemistry is central to water quality and water treatment modeling, but there are no standard open source water chemistry models. The most widely used model for water speciation chemistry, MINEQL+, is commercial and closed source. The development of desalination technologies and other water treatment approaches would benefit greatly from an open source tool.

Water System Analytics Tools. Water distribution and wastewater collection systems rely heavily on digital controls and contain vast amounts of information. Unwarranted access to these systems represents significant security risks. Potential issues include contamination of the supply, uncontrolled releases of critical supplies, flooding, etc. Big data approaches may help identify approaches to increase security and improve the efficiency of such systems.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Cyberinfrastructure has the potential to revolutionize the environmental and water resources engineering profession. Some key areas of need include advanced algorithms to simulate the behavior of water systems, open source computing tools, a computationally proficient workforce, and consistent, robust datasets. These tools are needed to solve emerging challenges in environmental and water resources management. New algorithms are needed to forecast streamflow and water quality, apply computational chemistry to water treatment processes, utilize bioinformatics to understand degradation of contaminants, and incorporate big data to enhance the sustainability and security of water systems. A culture of open source software development is needed to facilitate knowledge transfer, address challenges in maintenance of legacy codes, and increase the reproducibility of studies. An educational workforce must also be trained to incorporate computational thinking into the curriculum.

Hydrologic Modeling. Advanced cyberinfrastructure is needed to develop reproducible hydrologic models that do not require calibration. The critical drivers of these models are climate, land use, and economic development activities. Consistent, sustainable, publicly accessible datasets of key hydrologic observations are needed to develop such tools. New open source software constructed in high level programming languages may assist with automation of work flows and other essential aspects of a reproducible hydrologic modeling approach [8], [9].

Open Life Cycle Assessment Databases. LCA is an important tool for sustainability analysis. A publicly accessible, robust, geospatially and temporally explicit database of mass and energy flows for key economic processes is needed. Additional cyberinfrastructure would also need to be developed to calculate and analyze the environmental impact of product supply chains. Such tools could be extremely important in the development of science-based energy and environmental policies.

Open Source Water Treatment Process Models. Wastewater treatment plants are often very inefficient and consume large amounts of energy. Wastewater treatment plants contain many complex biochemical and physical processes. An open source modeling platform is needed to facilitate the optimization of these facilities. Inorganic water speciation chemistry is central to water quality and water treatment modeling. An open source water chemistry modeling platform should be developed. Such a tool would be of value for analysis of water and wastewater systems and for improved performance of desalination processes.

Water System Analytics Tools. Sustainable and secure water distribution and wastewater collection systems are essential for economic well-being. Because water systems rely heavily on digital controls and contain vast amounts of information, unwarranted access represents significant security risks such as water supply contamination, uncontrolled releases of critical supplies, flooding, etc. Movement of water also consumes large amounts of energy. Big data analytics approaches should be developed to increase security and improve the efficiency of such systems.

References
[1] N. H. Crawford and R. K. Linsley, “Digital Simulation in Hydrology: Stanford Watershed Model IV,” Stanford University, Department of Civil Engineering, 39, 1966.
[2] N. H. Crawford and S. J. Burges, “History of the Stanford watershed model,” Water Resour. Impact, vol. 6, no. 2, pp. 3–5, 2004.
[3] C. Hutton, T. Wagener, J. Freer, D. Han, C. Duffy, and B. Arheimer, “Most computational hydrology is not reproducible, so is it really science?,” Water Resour. Res., vol. 52, no. 10, pp. 7548–7555, 2016.
[4] C. Hutton, T. Wagener, J. Freer, D. Han, C. Duffy, and B. Arheimer, “Reply to Comment on ‘Most computational hydrology is not reproducible, so is it really science?,’” Water Resour. Res., 2017.
[5] L. A. Melsen, P. J. Torfs, R. Uijlenhoet, and A. J. Teuling, “Comment on ‘Most computational hydrology is not reproducible, so is it really science?,’” Water Resour. Res., 2017.
[6] M. N. Fienen, “HESS Opinions: Repeatable research: what hydrologists can learn from the Duke cancer research scandal,” Hydrol. Earth Syst. Sci., vol. 20, no. 9, p. 3739, 2016.
[7] P. L. McCarty, J. Bae, and J. Kim, “Domestic Wastewater Treatment as a Net Energy Producer–Can This be Achieved?,” Environ. Sci. Technol., vol. 45, no. 17, pp. 7100–7106, Sep. 2011.
[8] D. J. Lampert, PyHSPF: Python extensions to the Hydrological Simulation Program in Fortran. 2016.
[9] D. J. Lampert and M. Wu, “Development of an open-source software package for watershed modeling with the Hydrological Simulation

Program in Fortran," Environ. Model. Softw., vol. 68, pp. 166–174, Jun. 2015.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

The environmental and water resources engineering community stands to benefit tremendously from the development of novel cyberinfrastructure investments. Because of the current shortages in the application of advanced cyberinfrastructure, significant educational and workforce development will be needed. A culture of sharing and open source software and data must be fostered. Such changes in mentality could require substantial efforts to implement. Longevity and maintenance of cyberinfrastructure investments will be critical to the attainment of long-term goals.